

استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية

د. يحيى حياتي نصار

قسم علم النفس التربوي

الجامعة الهاشمية-الأردن

استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية

د. يحيى حياتي نصار

قسم علم النفس التربوي

جامعة الهاشمية - الأردن

الملخص

هدفت هذه الدراسة إلى تقديم مفهوم "حجم الأثر" أسلوباً إحصائياً مكملاً لفحص الفرضيات الإحصائية. تضمنت الدراسة تحليلاً رقماً لبرهنة أن حجم الأثر أقل تأثيراً بحجم العينة من الأساليب الإحصائية التي تستخدم لفحص الفرضيات. بالإضافة إلى ذلك قدمت الدراسة بعض المؤشرات الإحصائية التي تستخدم للدلالة على قيمة حجم الأثر في حالة استخدام بعض الأساليب الإحصائية. وبالتالي إختبار "ت" سواء في حالة العينات المترابطة أو المستقلة واختبار "ف" في حالة تحليل التباين الأحادي، ومعامل ارتباط بيرسون، وأخيراً الانحدار البسيط. كذلك حاولت الدراسة توضيح العلاقة بين حجم الأثر وقوة الاختبار الإحصائي حيث بينت أن حجم الأثر يرتبط إيجابياً مع قوة الاختبار المتوقع.

Effect Size as a Complementaly Statistical Procedure of Testing Hypotheses

Dr. Yahya H. Nassar

Dprt of Educational Pyschology
Hashemite University -Jordan

Abstract

This study aimed at presenting the concept of “effect size” as a complementary statistical procedure of testing hypotheses. The study included numerical analysis to prove that the magnitude of effect size is less influenced by the sample size than the statistical procedures used in testing hypotheses. Furthermore? the study displayed some statistical measures used to show the magnitude of effect size in the case of some statistical procedures. Specifically, t- test either in the case of independent samples or related samples, F- test in the case of univarite statistics? Pearson correlation coefficient? and simple regression. Also the study attempted to clarify the relationship between effect size and statistical power and indicated to the positive relationship between both of them.

استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية

د. يحيى حياتي نصار

قسم علم النفس التربوي

جامعة الهاشمية - الأردن

المقدمة :

منذ عدّة سنوات والباحثون ينصحون باستخدام حجم الأثر في تفسير نتائج أبحاثهم؛ وذلك للتمييز بين الدلالة العملية والدلالة الإحصائية لتلك النتائج (Mahadevan, 2000). ويشكك العديد من المتخصصين في منهجة البحث النفسي والتربوي، وفي العلوم الإنسانية عموماً، بالفائدة العملية للفرضيات الإحصائية في الإجابة عن كثير من الأسئلة البحثية (McClain, 1995). وبالتالي فإنّه من المعروف أن الفرضيات الإحصائية تجيب عن سؤال يتعلق بكون البيانات التي جمعها الباحث كافية لرفض الفرضية الصفرية (دالة إحصائية) أو غير كافية لرفض تلك الفرضية الصفرية (غير دالة إحصائية)، معنى إن الجواب عن الأسئلة الإحصائية يكون على الصورة: نعم توجد دلالة إحصائية أو لا توجد دلالة إحصائية (Thompson, 1998).

لكن السؤال الذي يبقى دائماً بحاجة إلى إجابة هو: إذا كانت الفروق دالة إحصائية أو غير دالة إحصائية، فإلى أي درجة كانت هذه الدلالة ذات فائدة عملية (Significant) وكذلك إلى أي درجة تشير عدم الدلالة الإحصائية إلى فروق ليست دالة على المستوى العملي (McLean & Ernest, 1997).

فعلى سبيل المثال قد يختبر باحث ما وجود فروق دالة إحصائية بين طرفيتين إرشاديتين في تخفيف قلق الامتحان لدى الطلبة، فإذا وصل ذلك الباحث إلى نتائج دالة إحصائية بحيث أشار إلى وجود فروق بين متوسطات المجموعات التجريبية والضابطة فإن السؤال الذي يجب أن يسأل لاحقاً هو إلى أي درجة كانت هذه الدلالة الإحصائية دالة من الناحية العملية؟، معنى إنه قد يكون الفرق بين متوسطي المجموعتين التجريبية والضابطة يساوي (٨) أو (١٥) أو (٢٠) أو (٣٠). وفي كل الأحوال قد يكون دالاً إحصائياً «لكن من الواضح أن فرقاً مقداره (٣٠) على سبيل المثال سيكون أكثر دلالة ليس فقط من ناحية رفض الفرضية الصفرية والحصول على نتائج دالة بل أيضاً من الناحية العملية من فرق مقداره (٨).

باختصار يمكن الحصول على فروق دالة إحصائية بوجوب فحص الفرضيات الإحصائية لمتوسطات متقاربة القيم فقط لأنّ الباحث قد اختار عينة كبيرة في دراسته، ومن المعروف

أن قيمة الخطأ المعياري (Standard Error) تقل كلما كبر حجم العينة، وبالتالي تكون الدلالة الإحصائية دالة لحجم العينة، أو لمستوى الدلالة بدلًا من أن تكون دالة للفروق العملية (practical differences) بين المتغيرات.

من هنا جاء الإهتمام بموضوع حجم الأثر (Size Effect)، حيث إن فكرة حجم الأثر تقوم بكل بساطة على صياغة الفروق بين المتوسطات وذلك باستخدام الانحراف المعياري (Standard Deviation) كوحدة قياس لمقدار الفرق بين تلك المتوسطات (كما هو الحال في اختبار (t) للعينات المترابطة والمستقلة) (Cohen, 1988) أو التعبير عن العلاقة بين المتغير المستقل أو المتغيرات المستقلة من جهة، والمتغير التابع أو المتغيرات التابعة من جهة أخرى، عن طريق استخراج حجم تباين المتغير التابع أو المتغيرات التابعة الذي يمكن تفسيره عن طريق المتغير المستقل، أو المتغيرات المستقلة، كما هو الحال في تحليل التباين (McClain, 1995)، أي أن قيمة حجم الأثر تبين إلى أي درجة يمكن التنبؤ بالمتغير التابع، أو تفسيره من خلال المتغير المستقل (Mahadevan, 2000).

ومن وجهة نظر ماكسويل وديلاني (Maxwell & Delaney, 1990) يمكن تقسيم المؤشرات التي تدل على مقدار حجم الأثر في الدراسات المختلفة إلى فئتين رئيسيتين هما:

١- المؤشرات التي تدل على مقدار الأثر الذي يحدده المتغير المستقل في المتغير التابع. وتعرف هذه المؤشرات أيضًا بأنها الفرق بين المتوسطات (متوسط المجموعة التجريبية ومتوسط المجموعة الضابطة في الدراسات التجريبية) بوحدات معيارية (عودة والخليلي، ١٩٨٨). وتستخدم هذه المؤشرات في الدراسات التي تعتمد على الأساليب الإحصائية التي تفحص الفروق بين متغيرات كما هو الحال في اختبار "t" سواء للعينات المستقلة أم العينات المترابطة. وأشهر هذه المؤشرات هو المؤشر الذي قدمه كوهين (1988, Cohen) ويطلق عليه الرمز "d" (وسيطلق عليه في الدراسة الحالية الرمز "د") ومؤشر جاما (γ) ومؤشر دلتا (Δ).

٢) المؤشرات التي تدل على قوة العلاقة أو الارتباط بين متغيرات الدراسة. وتستخدم هذه المؤشرات في الأساليب الإحصائية التي تعتمد على فحص التباين حيث تشير إلى نسبة التباين الذي يمكن تفسيره في المتغير التابع عن طريق تباين المتغير أو المتغيرات المستقلة، ومن أشهر هذه المؤشرات مؤشر مربع إيتا (Eta Squared) وقدم هذا المؤشر العالم فيشر (Fisher) عام ١٩٣٢، ومؤشر مربع أوميجا (Omega Squared) والذي قدمه هايز (Hays) عام ١٩٦٣، ومؤشر مربع إبسيلون (Epsilon Squared) والذي قدمه كيلي (Kelly) عام ١٩٣٥، ومؤشر مربع (ر) (R squared) والذي قدمه ستيفينز (Stevens) عام ١٩٩٢.

هناك العديد من الدراسات الأجنبية التي أشارت إلى أهمية استخدام حجم الأثر في الدراسات النفسية والتربوية والعلوم الإنسانية عموماً، ومن هذه الدراسات دراسة سنайдر

و لاوسن (Snyder & Lawson, 1993) حيث أجرى الباحثان دراسة هدفت إلى توضيح علاقة حجم العينة بقيمة الإحصائي، وبالتالي باحتمالية الحصول على نتائج دالة إحصائياً في حالة إفتراض ثبات قيمة حجم الأثر. وفي هذه الدراسة قدم الباحثان مثلاً تم فيه استخدام أسلوب تحليل التباين وبالتحديد الإحصائي (ف)، بحيث تم إجراء هذا التحليل لأكثر من مرة مع استخدام حجم عينة مختلف في كل تحليل ومع افتراض ثبات قيمة حجم الأثر حيث تم تثبيته (أي تثبت نسبة التباين المفسر إلى التباين الكلي) عند المستوى (٣٣, ٠, ٠)، وقد أوضحت هذه الدراسة أنه من الممكن الحصول على نتائج دالة إحصائياً من خلال زيادة حجم العينة على الرغم من ثبات قيمة حجم الأثر، وقد لاحظ الباحثان من خلال دراستهما أن القرار المتخذ حول رفض الفرضية الصفرية يمكن أن يتأثر وبصورة واضحة بحجم العينة ، وهذا يؤكد التساؤل حول دقة القرار المتخذ عند رفض الفرضية الصفرية في الدراسات التي تعتمد على أحجام عينات كبيرة. ويعني آخر فإنه في ضوء المثال المقدم في الجدول السابق فإنه يمكن للباحث عند اعتقاده بوجود علاقة ضعيفة بين متغيرات دراسته (المستقلة و التابعه) أن يزيد حجم عينة دراسته للحصول على نتائج دالة إحصائياً. وفي دراسة أخرى أشار كل من مكلين ويرنست (McLean & Eronest, 1997) إلى بعض نقاط الضعف في الاعتماد فقط على فحص الفرضيات الإحصائية في الأبحاث ، موضحين أنه يجب أن يشار إلى قيمة حجم الأثر بحسبه نتيجة مكملة لنتائج فحص الفرضيات ومعياراً آخر لفهم النتائج الإحصائية. وحسب رأي الباحثين فإنه في حين تفحص الفرضيات الإحصائية احتمال الحصول على نتائج دالة بمحض الصدفة فإن حجم الأثر يفحص مقدار الدلالة العملية للنتائج .

وتسائل ثومبسو (Thompson, 1998) في إحدى دراساته عن سبب مقاومة بعض الباحثين للإشارة إلى مقدار حجم الأثر في دراساتهم وإصرارهم على عدم تغيير الأساليب التقليدية في عرض النتائج الإحصائية في البحوث النفسية والتربوية. وقدم ثومبسو في هذه الدراسة العديد من النقاط الإيجابية التي يمكن تحصيلها من خلال استخدام حجم الأثر في الدراسات النفسية والتربوية.

وفي نفس السياق أشارت لين (Lane, 1999) إلى أن العديد من الدراسات التي أجريت في السنوات السابقة كانت تسيء تفسير النتائج الدالة لفحص الفرضية الصفرية، وأن هناك العديد من الباحثين الذين ضخمو من نتائج فحص الفرضيات الإحصائية المعتمدة فقط على رفض الفرضية الصفرية، وفي نفس الوقت قللوا من قيمة حجم الأثر أو أنهم غيبوا قيمته في الكثير من أبحاثهم. واستعرضت هذه الدراسة الأخطاء الشائعة في تفسير النتائج البحثية في ضوء فحص الفرضيات الصفرية.

أما دراسة فان (Fan, 2001) والتي اعتمدت على بيانات إفتراضية بإستخدام برامج حاسوبية معينة حيث يطلق على هذا النوع من الدراسات بدراسات مونت كارلو (Mont Carlo Studies) فقد أشارت هذه الدراسة إلى وجود علاقة بين فحص الدلالة

الإحصائية للبيانات ومقدار حجم الأثر وأنه يمكن وصف العلاقة بين الأسلوبين بالتكاملية، لكنها من جهة ثانية أشارت إلى أن أيًا من هذين الأسلوبين لا يعني عن الآخر، أو لا يمكن أن يكون بديلاً عن الآخر وانتهت هذه الدراسة إلى أن البحث الجيد يأخذ بعين الاعتبار الإشارة إلى نتيجة فحص الدلالة الإحصائية وقيمة حجم الأثر.

أما دراسة مكميلان ولوسون، ولوسون، وسوندر (Lawson, Mcmillan, & Synder, 2002) فقد هدفت بصورة أساسية إلى تنبية الباحثين حول الحاجة الملحة لفهم أهمية تضمين قيمة حجم الأثر في أبحاثهم. وتضمنت الدراسة شرحاً لمفهوم حجم الأثر وركزت على أهميته في الدراسات التربوية. ولفحص درجة اهتمام الأبحاث التربوية بإبراز قيمة حجم الأثر قام مكميلان وزملاؤه بمراجعة الأبحاث المنشورة في أربع مجالات تربوية متخصصة كعينة ممثلة للأدب التربوي. وتم التركيز في عملية المراجعة على الأبحاث الكمية، تم مراجعة (٥١٨) بحثاً صنفت كأبحاث كمية أو أبحاث خلطة (استخدمت المنهجين الكمي والنوعي في نفس الوقت) وقد وجدت الدراسة أن (١٤٨) بحثاً من هذه الأبحاث قد أشارت إلى قيمة حجم الأثر أو قامت بحسابه، ولكن (٨٢) دراسة فقط قامت بحسابه والاعتماد على قيمته لفحص الدلالة العملية للنتائج، و وجدت الدراسة أن (٣٠) دراسة فقط قامت بحسابه وحتوت على شرح تفصيلي ومكثف لمفهوم حجم الأثر وانتهت الدراسة إلى أن معظم الدراسات التي تم مراجعتها قد اعتمدت على فحص الدلالة الإحصائية في الوصول إلى نتائجها النهائية.

مشكلة الدراسة وأهميتها:

تكمن أهمية الإشارة إلى قيمة حجم الأثر في الدراسات النفسية والتربوية في أنها تمكّن الباحثين اللاحقين من استخدامها في الدراسات التي تستخدم أسلوب "التحليل التركيبي" (Meta Analysis studies) كما أن هذا الإجراء يزيد من إمكان الحكم على الدلالة العملية للنتائج وينسجم مع ما تضمنه دليل نشر جمعية علم النفس الأمريكية في أحد ث نسخه Mahadevan ، ٢٠٠٠ . وفي نفس السياق ومنذ عام ١٩٩٤ والجمعية النفسية الأمريكية (American Psychological Association) توّكّد على أهمية الإشارة إلى قيمة حجم الأثر في التقارير البحثية وذلك في ضوء تأثير نتائج فحص الدلالة الإحصائية للبيانات بحجم العينة، وفي عام ٢٠٠١ تم التأكيد على نفس النقطة من خلال الطبعة الخامسة من دليل نشر جمعية علم النفس الأمريكية (APA) لما ذلك من أهمية في توضيح النتائج البحثية. من هنا فإن العديد من المجالات البارزة في مجال علم النفس والعلوم التربوية أو العلوم الإنسانية تشرط على الباحثين تضمين قيمة حجم الأثر في أبحاثهم (Tapia & Marsh, 2002). ويفؤّد أرنيمان (Arnemann, 2003) على أن هناك (٢٣) مجلة متخصصة في مجال علم النفس التربوي والعلوم ذات العلاقة به – كالعلوم التربوية وعلم النفس – بما في ذلك مجلتان تعدان من أهم المجالات المتخصصة في هذا المجال

تطلب من الباحثين الراغبين في نشر الأبحاث الكمية من خلالها بتضمين أبحاثهم قيمة حجم الأثر.

وعلى الرغم من ذلك يلاحظ أن الكثير من الأبحاث النفسية والتربوية المنشورة في المجالات المتخصصة تخلو في كثير من الأحيان من الإشاره إلى قيمة حجم الأثر، ومن الغريب أن الباحث يبني دراسته على أساس مراجعه دقique للأدب يمكنه من إفتراض وجود فروق بين المتوسطات أو وجود علاقة بين المتغيرات المستقلة والمتغيرات التابعه مما يعني أنه افترض مسبق بوجود حجم أثر مقبول من المحتمل أن يمكنه من رفض الفرضية الصفرية الخاطئة وهذا ما يسمى بقوة الاختبار المتوقعة أي احتمال رفض الفرضية الصفرية الخاطئة، وإن هذا الاحتمال مرتبط منطقيا وإحصائيا بحجم الأثر المتوقع، لذلك إذا بني الباحث فرضياته على ضوء الدراسات السابقة التي تشير إلى وجود فروق دالة فضلاً عن إشارتها إلى حجم الأثر، فإنه بالتأكيد سيوفر كثيراً من جهده ووقته وربما من ماله أو مال مؤسسات متخصصة في دعم الأبحاث.

إن الهدف من البحث الحالي هو دعم فكرة استخراج حجم الأثر في الدراسات لما لذلك من فائدـة كبيرة، ولــما كان الكـثير من الباحثـين غير المتخصصـين في مجال منهـجـية الـبحث والإـحـصـاء النفـسي والتـربـوي غـير مـلمـين بماـهـيـة حـجمـ الأـثـرـ مـفـهـومـاـ وأـسـلـوبـاـ إـحـصـائـياـ وبـسبـبـ اـفـقارـ الأـدـبـ النفـسيـ والتـربـويـ عـلـىـ الصـعـيدـ العـرـبـيـ - حـسـبـ حدـودـ مـعـرـفـةـ الـبـاحـثـ - لـتـغـطـيـةـ شـامـلـةـ وـوـافـيـةـ لـهـذـاـ مـوـضـعـ مـنـ أـبعـادـ مـخـلـفـةـ ،ـفـإـنـ الـبـاحـثـ الـحـالـيـ يـحاـوـلـ أـنـ يـعـرـفـ الـبـاحـثـ العـرـبـيـ غـيرـ المـتـصـصـصـ فـيـ مـجـالـ الـبـحـثـ أـوـ إـحـصـاءـ.ـمـوـضـعـ حـجمـ الأـثـرـ مـفـهـومـاـ وأـسـلـوبـاـ إـحـصـائـياـ عـلـىـ أـمـلـ أـنـ يـسـهـمـ هـذـاـ الـبـحـثـ فـيـ تـشـجـيعـ الـمـجـالـاتـ الـعـرـبـيـهـ الـمـتـصـصـصـةـ عـلـىـ الـطـلـبـ مـنـ الـبـاحـثـينـ الإـشـارـةـ إـلـىـ قـيمـةـ حـجمـ الأـثـرـ فـيـ أـبـحـاثـهـ الـمـقـدـمةـ لـلـنـشـرـ مـنـ جـهـةـ وـعـلـىـ أـنـ يـقـتـنـعـ الـبـاحـثـونـ العـرـبـ بـأـهـمـيـةـ حـجمـ الأـثـرـ أـسـلـوبـاـ إـحـصـائـياـ مـكـمـلـاـ لـلـتـنـتـاجـ الـتـيـ يـتـمـ الـحـصـولـ عـلـيـهـاـ مـنـ خـالـلـ فـحـصـ الـفـرـضـيـاتـ إـحـصـائـيـةـ مـنـ جـهـةـ أـخـرىـ.

أسئلة الدراسة :

ستحاول الدراسة الحالية الإجابة عن الأسئلة التالية:

- ١- كيف يقلص حجم الأثر دور حجم العينة في فحص الفروق بين المتوسطات أو العلاقة بين المتغيرات المستقلة والتابعة؟
- ٢- كيف يتم استخراج حجم الأثر في الأساليب الإحصائية التي تعتمد على فحص الفروق بين المتوسطات؟
- ٣- كيف يتم استخراج حجم الأثر في الأساليب الإحصائية التي تعتمد على تحليل التباين؟
- ٤- كيف يتم استخراج حجم الأثر في الدراسات الارتباطية (correlational)؟
- ٥- ماهي معايير الحكم على حجم الأثر (صغير أو متوسط أو كبير) في الأساليب

الإحصائية السابقة؟

٦- ما هي علاقة مقدار حجم الأثر بقوة الاختبار المتوقعة؟

٧- ما هي علاقة مقدار حجم الأثر بتحديد حجم العينة المناسب في الدراسات المختلفة؟

منهجية الدراسة واجراءاتها:

بهدف الإجابة عن أسئلة الدراسة الحالية سيتم بدايةً التعريف بمفهوم "حجم الأثر" بعرض توضيحه وربطه بالمفاهيم الإحصائية الأخرى وخاصةً الفرضيات الإحصائية. ثم سيتم توضيح العلاقة بين حجم عينة الدراسة من جهة والدلالة الإحصائية وحجم الأثر من جهة أخرى، وفي هذا السياق سيتم تقديم مثال افتراضي باستخدام الإصدار العاشر من الرزمة الإحصائية للعلوم الإنسانية (SPSS) وبالتحديد سيتم من خلال المثال الافتراضي توضيح العلاقة بين حجم العينة والدلالة الإحصائية عند ثبيت مقدار ز حجم الأثـرس، وفي الخطوة التالية سيتم عرض بعض مؤشرات "حجم الأثر" التي تستخدم في الأساليب الإحصائية المختلفة، ثم سيتم عرض العلاقة بين "حجم الأثر" و "قوة الاختبار المتوقعة" ، وأخيراً سيتم إستعراض العلاقة بين "حجم الأثر" وحجم العينة المناسب في الدراسات المختلفة.

مفهوم "حجم الأثر":

لقد قامـت محاولات الباحثين في تفسير مفهوم "حجم الأثر" بدور هام في زيادة المعرفة الإحصائية والبحثية في مجال العلوم الإنسانية، وعلى كل حال هناك بعض الغموض في تحديد وتفسير معنى قيمة حجم الأثر (Fern & Monroe, 1996) . ومن المهم بدايةً تحديد المقصود بحجم الأثر. فمن وجهة نظر كوهين (Cohen, 1977) فإنه في الوقت الذي تعبـر فيه الفرضية الصفرية عن غياب الظاهرة قـيد الدراسة في مجتمع ما (كأن نقول أنه لا تـوجد عـلاقة دـالة إحـصـائـيـاً بين مـسـتوـى دـافـعـيـة طـلـبـةـ الـثانـوـيـةـ وـتحـصـيلـهـمـ الأـكـادـيـيـيـ وـبـحـيثـ يتم استـخدـامـ قـيمـةـ إـحـصـائـيـةـ مـحدـدـةـ كـعـامـلـ اـرـتـباطـ يـبـرـسـونـ عـلـىـ سـبـيلـ المـثالـ لـتـعـبـيرـ عـنـ هـذـاـ المعـنىـ)، فإن حـجمـ الأـثـرـ يـحـمـلـ عـكـسـ هـذـاـ المعـنىـ حيث عـرـفـهـ كـوهـينـ (Cohen, 1997) على أنه "درجة توـاجـدـ ظـاهـرـةـ فـيـ مجـتمـعـ ماـ (ـفـروـقـ بـيـنـ مـتوـسـطـاتـ أوـ عـلـاقـةـ بـيـنـ مـتـغـيرـاتـ فـيـ مجـتمـعـ مـحـدـدـ)ـ أوـ هوـ درـجـةـ خـطـأـ الفـرـضـيـةـ الصـفـرـيـةـ (ـأـيـ درـجـةـ وـجـودـ فـرقـ بـيـنـ المـتوـسـطـاتـ أوـ عـلـاقـةـ بـيـنـ المـتـغـيرـاتـ)"ـ (ـصـ ٩ـ).ـ وهذاـ يـعـنيـ أنهـ عـنـدـماـ تـكـوـنـ الفـرـضـيـةـ الصـفـرـيـةـ خـاطـئـةـ فإـنـ هـذـاـ لاـ يـعـنـيـ بـالـضـرـورـةـ وـجـودـ فـرقـ كـبـيرـ وـوـاضـحـ بـيـنـ المـتوـسـطـاتـ ،ـ حيثـ يـمـكـنـ أـنـ يـكـونـ هـذـاـ فـرقـ صـغـيرـأـ أوـ مـتوـسطـأـ ،ـ منـ هـنـاـ يـمـكـنـ اـحـتـسابـ حـجمـ الأـثـرـ قـيمـةـ أـكـبـرـ مـوـجـودـةـ فـيـ مجـتمـعـ مـحـدـدـ وـأـنـهـ كـلـمـاـ كـبـرـتـ هـذـهـ الـقـيمـةـ كـلـمـاـ زـادـتـ اـحـتمـالـيـةـ توـاجـدـ الـظـاهـرـةـ قـيدـ الـدـرـاسـةـ فـيـ ذـلـكـ الـجـمـعـ.ـ إـنـهـ وـفـيـ ضـوءـ التـعـرـيفـ السـابـقـ لـاـ بـدـ مـنـ التـميـزـ بـيـنـ الـعـدـيدـ مـنـ الـمـصـطـلـحـاتـ الـتـيـ تـسـتـخـدـمـ فـيـ الإـشـارـهـ إـلـىـ مـقـدـارـ حـجمـ الأـثـرـ.ـ بـالـنـسـبـهـ لـقـيمـةـ حـجمـ الأـثـرـ فـيـ

المجتمع (مجتمع الدراسة) هي قيمة أكبر من الصفر وهي القيمة الحقيقية غير المعروفة للأثر الممثل للعلاقة بين المتغيرات المراد دراستها. فإذا لم توجد علاقة بين متغيرات الدراسة في المجتمع فإن حجم الأثر لهذه العلاقة سيكون صفرًا أو قريباً من الصفر. أما إذا كان هناك علاقة بين متغيرات الدراسة ضمن المجتمع المحدد فإنه يمكن تقدير هذه العلاقة من خلال عينة ممثلة لذلك المجتمع (لها نفس خصائص مجتمع الدراسة).

أما بالنسبة لمؤشر حجم الأثر (Effect Size Indicator) فهو قيمة كمية تشير إلى درجة العلاقة بين متغيرات الدراسة ضمن مجتمع محدد مسبقاً (مجتمع الدراسة). وفي الدراسة الحالية سيتم استخدام مصطلح حجم الأثر (Effect Size) للإشارة إلى قيمة حجم الأثر المقدرة من خلال عينة ممثلة لمجتمع محدد وبحيث تشير هذه القيمة إلى مقدار ذلك الأثر في ذلك المجتمع بغض النظر عن الأسلوب الإحصائي المستخدم للتعبير عن ذلك الأثر.

أثر حجم عينة الدراسة في الدلالة الإحصائية وحجم الأثر:

إن من الميزات المهمة لمقدار حجم الأثر أنه يمكن الباحث في أي دراسة من تقدير مدى تأثير الدلالة الإحصائية للنتائج أي بحجم العينة عدد الوحدات المختارة من ضمن مجتمع معين لتمثيل ذلك المجتمع (Snyder & Lawson, 1993). على الرغم من وجود العديد من العوامل المتداخلة التي تؤثر في الحصول على نتائج دالة إحصائية (Schneider & Darcy, 1984)، يعد حجم العينة العامل الأساس في تقرير ما إذا كانت نتائج أي دراسة دالة أو غير دالة إحصائية (Snyder & Lawson, 1993). وكما يشير كريج وآيسون وميتز (Craig, Eison & Metz, 1976) إنه كلما كبر حجم عينة الدراسة كلما زاد احتمال الحصول على نتائج دالة إحصائية حتى ولو كان هناك ارتباط ضعيف ضمن مجتمع الدراسة بين المتغيرات المستقلة والمتغيرات التابعة. وقد لاحظ هايز (Hays, 1963) بأن الدلالة الإحصائية لا تشير بالضرورة إلى قوة العلاقة بين متغيرات الدراسة، بل هي في أحسن الأحوال تشير إلى بعض الارتباط بين هذه المتغيرات، دون تحديد مدى أهمية ذلك الارتباط، هذا وقد أضاف هايز (Hays, 1963) في هذاخصوص أن العامل الأهم في الحصول على نتائج دالة إحصائية هو حجم العينة، وأنه من المهم التحديد الدقيق لمفهوم قوة العلاقة الإحصائية، ويمكن تفسير ملاحظات هايز على أساس اعتبار حجم العينة مطلقاً في كثير من الأحيان، بحيث أن زيادة حجم العينة يؤدي إلى الحصول على نتائج دالة إحصائية، مما يضخم من حجم العلاقة بين المتغيرات المستقلة والمتغيرات التابعة وما يحد من القدرة على تحديد حجم العلاقة الحقيقة بين هذه المتغيرات. ولقد قدم كل من ماكسويل و ديلاني (Maxwel & Delaney, 1990) معادلة لتوضيح العلاقة بين قيمة الإحصائي (χ^2 ، t ، F ، كاي تربيع... الخ) وحجم عينة الدراسة وحجم الأثر، وفيما يلي عرض لتلك المعادلة:

قيمة الإحصائي = حجم الأثر \times حجم الدراسة
 يتضح من المعادلة السابقة أن قيمة الإحصائي المستخدم لفحص الفرضيات الإحصائية في أي بحث يعتمد على عاملين أساسين هما "حجم الأثر" أي قوة العلاقة بين متغيرات الدراسة و "حجم الدراسة" الذي ما هو إلا دالة لحجم العينة حسب رأي سنайдر ولوسن (Snyder & Lawson, 1993) ويشار إليه عادةً في الدراسات بدرجات الحرية. (Degrees of Freedom) ومن المعروف إحصائياً بأنه كلما دلت نتائج أي دراسة إلى قيمة إحصائية أكبر كلما زادت احتمالية رفض الفرضية أو الفرضيات الصفرية المتضمنة في تلك الدراسة، فعلى سبيل المثال إذا استخدم باحث إختبار (ت) لفحص فرضيات دراسته فإنه يستطيع الافتراض بأنه كلما كبرت قيمة (ت) المحسوبة من خلال بيانات الدراسة كلما زادت احتمالية الحصول على نتائج دالة إحصائية. ومن خلال المعادلة السابقة و يمكن للباحث في حالة افتراضه قيمة ثابتة لحجم الأثر أن يفترض بأن قيمة الإحصائي ما هي إلا دالة لحجم العينة. معنى أنه كلما كبر حجم العينة كلما زادت قيمة الإحصائي و زادت احتمالية الحصول على نتائج دالة إحصائية.

ولتأكيد علاقة حجم العينة باحتمالية رفض الفرضية الصفرية أو الفشل في رفضها عند تثبيت قيمة حجم الأثر تم افتراض بعض البيانات الرقمية من قبل الباحث والتي تمثل علامات مفترضة لمجموعة من الطلبة في مقرر معين . وقد تم افتراض أن الطلبة قد قسموا إلى ثلاث مجموعات بحيث تم تدريس المجموعة الأولى بوساطة الطريقة (أ) والمجموعة الثانية بالطريقة (ب) والمجموعة الثالثة بالطريقة (ج) ومن ثم تم تحليل هذه البيانات مع أحجام عينيات مختلفة باستخدام أسلوب تحليل التباين وبالتحديد الإحصائي (ف) على افتراض أن أفراد المجموعات الثلاث قد تم سحبها من مجتمعات تتوزع توزيعاً سوياً ومتجانسة من حيث التباين ، وقد تم تثبيت مقدار حجم الأثر عن طريق مضاعفة حجم العينة إلى عدد من المرات ولكن مع استخدام نفس البيانات التي تم استخدامها في أول تحليل وهكذا كان حجم العينة يتزايد عن طريق تكرار نفس البيانات ، ولكن مع حجم أثر ثابت. والمجدول رقم (١) يظهر البيانات التي تم استخدامها في أول تحليل.

المجدول رقم (١)

ويظهر البيانات التي تم استخدامها في التحليل الأول

رقم الطالب	٦	٥	٤	٣	٢	١
طريقة التدريس	ج	ج	ب	ب	أ	أ
العلامة	٦٦	٦٩	٧٠	٦٧	٦٨	٦٥

يتضح من خلال المجدول رقم (١) أنه قد تم البدء افتراضياً بستة طلاب بحيث ضمت كل

مجموعه طالبين مع الأخذ بعين الاعتبار أن أسلوب تحليل التباين يتطلب عدداً أكبر من الأفراد في كل مجموعه ولكن تم البدء بهذا العدد الصغير (فردان في كل مجموعه) فقط لإيضاح العلاقة بين حجم العينة واحتمالية رفض الفرضية الصفرية عند ثبيت مقدار حجم الأثر. هذا وقد تم استخدام البيانات الواردة في الجدول رقم (١)، ولكن بعد مضاعفة أعداد المجموعات الثلاث من مرة إلى ست مرات. فعلى سبيل المثال كانت البيانات في المرة الثانية بعد مضاعفة بيانات كل مجموعه لمرة واحدة على النحو التالي:

المجموعه الأولى: ٦٨، ٦٨، ٦٥، ٦٥

المجموعه الثانية: ٧٠، ٧٠، ٦٧، ٦٧

المجموعه الثالثة: ٦٦، ٦٦، ٦٩، ٦٩

وهذا تم إجراء تحليل التباين ست مرات، والجدول رقم (٢) يوضح نتائج هذه التحاليل.

الجدول رقم (٢)

نتائج تحليل التباين للبيانات المفترضة للعلاقة بين طريقة التدريس وتحصيل الطلبة في

مقرر معين عند تغيير حجم العينة (٣٦، ٣٠، ٢٤، ١٨، ١٢، ٦)

وتبثيت مقدار حجم الأثر (٠٠,٢٢٩)

مصدر التباين	مجموع مربع الإنحرافات	مربع إيانا (حجم الأثر)	درجات الحرية	متوسط مجموع مربع الإنحرافات	قيمة F	مستوى الدلالة الفعلية	القرار
طريقة التدريس	٤,٠٠	٠,٢٢٩	٢	٢,٠٠	٠,٤٤٤	٠,٦٧٨	عدم رفض الفرضية
	١٣,٥٠		٣	٤,٥٠			
	١٧,٥٠		٥				
	٦=حجم العينة						
خطأ	٨,٠٠	٠,٢٢٩	٢	٤,٠٠	١,٣٣	٠,٣١١	عدم رفض الفرضية
	٢٧,٠٠		٩	٣,٠٠			
	٣٥,٠٠		١١				
	١٢=حجم العينة						
طريقة التدريس	١٢,٠٠	٠,٢٢٩	٢	٦,٠٠	٢,٢٢	٠,١٤٣	عدم رفض الفرضية
	٤٠,٥٠		١٥	٢,٧٠			
	٥٢,٥٠		١٧				
	١٨=حجم العينة						
خطأ	١٦,٠٠	٠,٢٢٩	٢	٨,٠٠	٣,١١	٠,٠٦٦	عدم رفض الفرضية
	٥٤,٠٠		٢١	٢,٥٧			
	٧٠,٠٠		٢٣				
	٢٤=حجم العينة						

تابع الجدول رقم (٢)

مصدر التباين	مجموع مربع الإنحرافات	مربيع إيتا (حجم الأثر)	درجات الحرية	متوسط مجموع مربع الإنحرافات	قيمة ف	مستوى الدلالة الفعلية	القرار
طريقة التدريس	٢٠,٠٠	٠,٢٢٩	٢	١٠,٠٠	٤,٠٠	٠,٠٣٠	رفض الفرضية
	٦٧,٥٠		٢٧	٢,٥٠			
	٨٧,٥٠		٢٩				
	٣٠ = حجم العينة						
طريقة التدريس	٢٤,٠٠	٠,٢٢٩	٢	١٢,٠٠	٤,٨٨	٠,٠١٤	رفض الفرضية
	٨١,٠٠		٣٣	٢,٤٥			
	١٠٥,٠٠		٣٥				
	٣٦ = حجم العينة						

بالنظر إلى الجدول رقم (٢) يمكن استخلاص النقاط التالية:
أولاً: إن قيمة مربيع إيتا والتي تمثل مقدار التباين المفسر لدرجات الطلبة في المقرر المعين من خلال تباين متغير طريقة التدريس بقيمة ثابتة حيث بلغت في جميع التحليلات (٠٠,٢٢٩). وهذا يعني أن تباين متغير طريقة التدريس فسر (افتراضياً) ما نسبته تقريباً ٢٣٪ من تباين متغير درجات الطلبة في المقرر المقترض. وبتعبير آخر لم تتأثر قيمة حجم الأثر بتزايد حجم العينة في التحاليل الستة المختلفة حيث تزايد حجم العينة من (٦) إلى (٣٦) دون أن يؤثر ذلك على مقدار مربيع إيتا. ومربع إيتا هو أحد المؤشرات الإحصائية التي تستخدم لفحص مقدار التباين المفسر للمتغير التابع عن طريق تباين المتغير المستقل عند استخدام أسلوب تحليل التباين، وسيتم لاحقاً تقديم معلومات أوسع حول هذا المؤشر وبعض المؤشرات الأخرى التي تستخدم في الأساليب الإحصائية المختلفة.

ثانياً: أن قيمة الإحصائي (ف) ازدادت مع التزايد الافتراضي لحجم العينة حيث بلغت (٤٤,٠) عندما كان عدد أفراد العينة (٦) ثم بلغت (١,٣٣) عندما ازداد حجم العينة إلى (١٢) فرداً ثم أصبحت (٢,٢٢) عندما ازداد حجم العينة إلى (١٨) فرداً كما بلغت قيمة (ف) (١١,١٣) عند وصول عدد أفراد العينة إلى (٢٤). ثم ازدادت قيمة (ف) عندما ازداد حجم العينة إلى (٣٠). وأخيراً بلغت (٤,٨٨) عندما وصل حجم العينة إلى (٣٦). وهذا يؤكد دقة المعادلة التي قدمها كل من ماكسويل و ديلاني (Delaney, 1990) والتي تشير إلى إمكان زيادة قيمة الإحصائي عند تثبيت مقدار حجم الأثر عن طريق زيادة حجم العينة.

ثالثاً: مع تزايد قيم (ف) المطرد مع ازدياد حجم العينة فإنه من الطبيعي ازدياد احتمالية رفض

الفرضية الصفرية. ومن الواضح أنه على الرغم من ثبات مقدار قيمة حجم الأثر (٠,٢٢٩) أصبح من الممكن رفض الفرضية الصفرية والتي تشير إلى عدم وجود فروق دالة إحصائياً بين متوسطات المجموعات الثلاث من الطلبة الذين تم تدريفهم بالطرق المفترضة الثلاث (أ، ب، ج) حيث بلغ حجم العينة (٣٠) فرداً (عشرة أفراد في كل مجموعة). وهذا يؤكد أن إحتمال رفض الفرضية الصفرية يتزايد مع تزايد حجم العينة، وذلك ليس كمؤشر لقوة العلاقة بين المتغير المستقل أو المتغيرات المستقلة والمتغير التابع أو المتغيرات التابعه ولكنه نتيجة لكبر حجم العينة.

المؤشرات الإحصائية المستخدمة للدلالة على حجم الأثر في الأساليب الإحصائية المختلفة :

في ما يلي استعراض لبعض أنواع المؤشرات التي تستخدم في الأساليب الإحصائية وبالتحديد في اختبار "ت" (t-test) للعينات المستقلة والمترابطة، واختبار تحليل التباين الاحادي (Pearson-momen Coefficient) ومعامل ارتباط بيرسون (ANOVA) وتحليل الانحدار (Regression Analysis).

١ المؤشرات التي تستخدم للدلالة على حجم الأثر في اختبار "ت" للعينات المترابطة والمستقلة:

أ- مؤشر كوهين "د" (Cohen's d) : تقوم فكرة هذا المؤشر الذي قدمه كوهين (Cohen,1998) على استخراج الفرق بين متوسط المجموعة التجريبية والمجموعة الضابطة ومن ثم قسمة هذا الفرق على الانحراف المعياري لإحدى المجموعتين على أن كلا المجموعتين متجلانستين من حيث التباين، وهذا هو أحد الافتراضات التي يفضل توفرها عند إستخدام اختبار "ت" بالإضافة إلى افتراض استقلالية البيانات وافتراض التوزيع الطبيعي للبيانات في مجتمع أو مجتمعات الدراسة وهذه الافتراضات يجب التتحقق منها قبل استخدام أي أسلوب إحصائي معلمي (parametric statistical procedure) ويمكن عرض المعادلة التي قدمها كوهين لاستخراج حجم الأثر في حالة إختبار "ت" على النحو التالي:

$$d = \frac{M_t - M_p}{s}$$

ع

حيث إن (M_t) : متوسط المجموعة التجريبية ، (M_p) : متوسط المجموعة الضابطة، (s) : الانحراف المعياري لأي من المجموعتين

فعلى سبيل المثال لو كان متوسط المجموعة التجريبية يساوي (٢٠٨) ومتوسط المجموعه الضابطة يساوي (٢٠٠) والانحراف المعياري للمجموعة التجريبية يساوي (٤٨) والانحراف المعياري للمجموعة الضابطة يساوي (٤٥) وأن المجموعتين متجلانستين من

حيث التباين فإنه يمكن حساب حجم الأثر حسب مؤشر كوهين "د" على النحو التالي:

$$d = \frac{200-208}{48} = \frac{200-208}{48} = 0,17$$

وهذا يعني أن الفرق بين متوسط المجموعة التجريبية والمجموعة الضابطة هو (١٦,٠) أو (١٧,٠) بوحدات الانحراف المعياري.

ومن المعروف أنه عندما تكون العينتين التجريبية والضابطة غير متساويتين في الحجم (عدد أفراد إحدى المجموعتين أكبر من الأخرى) فإن هناك احتمال كبير لأن يتحقق إفتراض تساوي التباين بين المجموعتين لذلك يفضل في حالة عدم تساوي تباين المجموعتين قسمة الفرق الناتج عن اختلاف قيمة المتوسطات على الجذر التربيعي لما يسمى في حالة إختبار "ت" بالتباین الممزوج (pooled variance) والذي هو عبارة عن دمج تباين كلا المجموعتين الممثلتين بمحتملين مختلفين بتوزيعين مستقلين في توزيع واحد، وبالتالي واحد ويعبر عن هذا التباين بالمعادلة التالية:

$$d^2 = \frac{(n_1-1) s_1^2 + (n_2-1) s_2^2}{n_1+n_2-2}$$

حيث إن (n_1) : هو عدد أفراد المجموعة الأولى ، (n_2) : هو عدد أفراد المجموعة الثانية (s_1^2) : هو تباين المجموعة الأولى ، (s_2^2) : هو تباين المجموعة الثانية (d) : هو التباين الممزوج.

فعلى سبيل المثال لو افترضنا أن التباين الممزوج لكلا المجتمعين يساوي (٢٠٩) فإنه في هذه الحالة يمكن استخراج قيمة المؤشر "د" على النحو التالي:

$$d = \frac{200-208}{\sqrt{209}} = 0,17$$

أما في حالة إختبار "ت" للعينات المرابطة فإنه عند استخدام مؤشر "د" لتقدير قيمة حجم الأثر فإنه يتم استخدام نفس القانون ولكن باستخدام المتوسطات القبلية والبعديه لنفس المجموعة (في حالة إجراء قياس قبلي وقياس بعدى لنفس المجموعة) لاستخراج الفرق وبالقسمة على الانحراف المعياري لأحد القياسيين (قبلى أو بعدى) في حالة الوفاء بافتراض تجانس التباين، أو بالقسمة على الانحراف المعياري للفروق في حالة استخدام الفروق بين القياس القبلي والقياس البعدي لاستخراج قيمة "ت" حيث تصبح معادلة مؤشر حجم الأثر "د" في هذه الحالة كما يلي:

$$d = \frac{m_c - m_b}{s}$$

حيث إن (م ق): متوسط المجموعة القبلي، (م ب): متوسط المجموعة البعدى، (ع ع): الانحراف المعياري للمجموعة في القياس القبلي أو البعدى

$$\text{أو } d = \frac{M_f - M_c}{S}$$

حيث إن (م ف): متوسط الفروق بين القياس القبلي والقياس البعدى، (ع ف): الانحراف المعياري للفروق.

ومن الجدير ذكره أنه عند التعامل مع مؤشر حجم الأثر "d" فإنه يجب تجاهل الإشارة السالبة أي أنه يجب التعامل مع القيمة المطلقة لهذا المؤشر (Howell, 1995). معنى أنه لو كان الفارق بين المجموعة التجريبية والمجموعة الضابطة قيمة سالبة فإنه بقسمة هذه القيمة على الانحراف المعياري للتباين الممزوج فستكون قيمة مؤشر "d" الناتجة في هذه الحالة قيمه سالبه ويعنى أن هذه القيمة تعبر عن عدد الوحدات المعيارية التي ينحرفها أحد المتوضطين عن المتوسط الآخر، وبما أن الانحراف المعياري هو بالضرورة قيمة موجبة فإنه يجب التعامل مع القيمة المطلقة لمؤشر "d" وبحيث يتم تجاهل الإشارة السالبة.

ومن المؤشرات الأخرى التي تستخدم لاستخراج حجم الأثر في حالة اختبار "t" أو فحص الفروق بين متوضطين مؤشر جاما (٧) وهو لا يختلف عن مؤشر كوهين "d" فهو يعتمد على استخراج الفرق بين المتوضطين ومن ثم قسمة الناتج على الانحراف المعياري للتباين الممزوج (Howell, 1995) وكذلك الحال بالنسبة لمؤشر دلتا (٨) الذي يعتمد على نفس المبدأ (عوده والخليلي، ١٩٨٨).

ولقد أشار كوهين (Cohen, 1988) إلى معايير للحكم على قيمة حجم الأثر المستخرج بواسطة مؤشر "d" حيث عده صغيراً عند القيمة (٠,٢٠) وعده متوسطاً عند القيمة (٠,٥٠) وعده كبيراً عند القيمة (٠,٨٠)، معنى أنه لو كانت قيمة مؤشر "d" والتي تعبّر عن الفرق بين متوسط تحصيل طلبة السنة الأولى في جامعة ما، وطلبة السنة الثانية في اختبار مستوى في اللغة العربية تساوي (٠,٢٣) فإن هذا يعني وجود فرق "صغير" بين المتوضطين وكأننا نشير من ناحية عملية إلى أن فارق سنة جامعية كاملة بين الفتتین لم يكن كافياً لإظهار فرق واضح، أو ظاهر بين طلبة السنة الأولى والسنة الثانية في اللغة العربية.

وأشار هاويل (Howell, 1995) إلى أنه يمكن اعتماد نفس المعايير عند استخدام جاما (٧) كمؤشر لحجم الأثر ويمكن القول أن نفس المعايير يمكن أن تستخدم للحكم على مقدار حجم الأثر عند استخدام مؤشر دلتا (٨). ومن الجدير ذكره أن المعايير السابقة قد وضعت على أساس دراسة قام بها كوهين حيث أثبت أنه عندما يكون حجم الأثر (٠,٢٠) فإن درجة التداخل (overlap) بين توزيع المجتمع الأول وتوزيع المجتمع الثاني (على أساس افتراض توزيعين بمحتمعين في حالة اختبار "t") تصل إلى ٨٥٪ وأن درجة التداخل بين

التوزيعين تصل إلى ٦٧٪ عندما تكون قيمة حجم الأثر ٥٠، بينما تصبح درجة التداخل ٥٣٪ عندما تبلغ قيمة حجم الأثر (Aron & Aron, 1997) (٨٠، ٥٠).

لكن كيف يقلص حجم الأثر وبالتحديد مؤشر كوهين "د" دور حجم العينة؟ من المعروف مثلاً أن الانحراف المعياري للتوزيع العيني للمتوسطات المستخدم في اختبار "ت" للعينة الواحدة ما هو إلا الخطأ المعياري المستخدم في حالة حساب قيمة "ت"، عند مقارنة متوسط عينة بمتوسط مجتمع معلوم مسبقاً دون معرفة تباين ذلك المجتمع حيث يتم تقديره من خلال العينة وفي هذه الحالة يتم حساب قيمة "ت" وفق المعادلة التالية:

$$t = \frac{\bar{m} - m}{\sqrt{\frac{s^2}{n}}}$$

حيث إن (\bar{m}) : متوسط العينة ، (m) : متوسط المجتمع ، (s) : الانحراف المعياري للعينة ، (n) : حجم العينة.

ومن خلال المعادلة السابقة يمكن ملاحظة أنه كلما كبر حجم العينة كلما قل الخطأ المعياري، (s) وكلما زادت قيمة "ت" ازدادت إحتمالية رفض الفرضية الصفرية. من هنا فإنه يمكن رفض الفرضية الصفرية على الرغم من أن الفرق بين متوسط العينة ومتوسط المجتمع قد يكون صغيراً لكن تضخيم حجم العينة هو الذي أدى إلى تضخيم قيمة "ت" وبالتالي تم رفض الفرضية الصفرية . لكن مؤشر حجم الأثر في مثل هذه الحاله والذي يتم حسابه وفق المعادله التاليه:

$$d = \frac{\bar{m} - m}{s}$$

يستبعد قيمة (n) أو حجم العينة من المعادلة ومن هنا يتم تقدير قيمة الفرق بين متوسط المجتمع ومتوسط العينة بوحدات معيارية بعد "تقليص" دور حجم العينة حيث يبقى تأثير حجم العينة وليس استبعاده، على قيمة الانحراف المعياري للعينة . وما ينطبق على اختبار "ت" للعينة الواحدة ينطبق على اختبار "ت" سواء في حالة العينات المستقلة أو المترابطة.

٢- المؤشرات التي تستخدم للدلالة على حجم الأثر في تحليل التباين الأحادي :
 تعتمد المؤشرات التي تستخدم للدلالة على حجم الأثر- في حالة استخدام تحليل التباين الأحادي لفحص فرضيات البحث- على رصد نسبة التباين في المتغير التابع والمرتبط بتباين المتغير أو المتغيرات المستقلة وتراوحاً قيمة هذه المؤشرات من صفر إلى ١ . ومن الأمثله على هذه المؤشرات كما يشير سنايدر ولوسن (Snyder & Lawson, 1993) مؤشر مربع إيتا (Eta Squared) ويرمز لها في الأبحاث الإحصائية المتخصصة

والمنشورة باللغة الإنجليزية بالرقم (Fesher) عام ١٩٣٢، ومؤشر مربع إيتا الجزئي (Partial Eta Squared) والذي قدمه كوهين (Cohen) عام ١٩٧٣، ومؤشر مربع أوميغا (Omega Squared) والذي قدمه هايز (Hays) عام ١٩٦٣، ومؤشر مربع إبسيلون (Epsilon Squared) والذي قدمه كيلي (Kelly) عام ١٩٣٥، ومؤشر مربع (R squared) والذي قدمه ستيفينز (Herzberg Formula) عام ١٩٩٢، ومعادلة هيرزبيرغ عام ١٩٦٩ (Stevens) والتي قدمها هيرزبيرغ عام ١٩٦٩، وأخيراً معادلة لورد (Lord Formula) والتي قدمها لورد عام ١٩٥٠.

وقد قدم الباحثون من خلال دراساتهم وأبحاثهم طرقاً مختلفة لحساب كل من المؤشرات السابقة وللروا على أنه يمكن استخدام هذه المؤشرات في الأبحاث ذات التصاميم التجريبية والأبحاث الارتباطية على حد سواء (Snyder & Lawson, 1993). هذا وقد أشار سنايدر ولوسن (Snyder & Lawson, 1993) إلى أن بعض المؤشرات السابقة تضخم من نسبة التباين المفسر، كما هو الحال مع مؤشر مربع إيتا ومؤشر مربع (r) وهذا يعني أن قيم هذه المؤشرات والمستخرجة من خلال العينات، هي قيم مضخمة بحيث أنه إذا تم استخراجها من خلال المجتمعات التي سُحبَت منها تلك العينات فإنها ستكون أقل من تلك التي تم تقديرها من خلال العينات.

ولعل أكثر المؤشرات شيوعاً لتقدير قيمة حجم الأثر في حالة استخدام تحليل التباين الأحادي هو مؤشر مربع إيتا ومؤشر مربع أوميغا وذلك لسهولة حسابه (Howell, 1995) وفيما يلي عرض لكلا المؤشرتين:

أ- مؤشر مربع إيتا (η^2): على الرغم من كون هذا المؤشر متحيزاً فإنه يضخم من مقدار الارتباط بين تباين المتغير التابع و تباين المتغير المستقل، والمقدر من خلال العينة مقارنة بالقيمة التي يتم استخراجها من خلال جميع أفراد مجتمع الدراسة، وتعد عملية حساب هذا المؤشر في غاية السهولة ومن السهل تفسير القيمة التي يتم الحصول عليها من خلالها. يتم الاستدلال عند استخدام أسلوب تحليل التباين الأحادي على مقدار التباين الكلي في بيانات الدراسة من خلال ما يعرف بمجموع مربع الانحرافات الكلي. ويعود جزء من هذا التباين الكلي في البيانات إلى أثر المعالجة حيث يفسر الاختلاف في درجات جمومعات الدراسة المختلفة إلى أثر هذه المعالجة ويمكن أن يعود جزء آخر من من التباين الكلي إلى متغيرات مستقلة أو ثانوية أخرى في الدراسة وإلى التفاعل ما بين هذه المتغيرات، بينما يعود باقي التباين إلى ما يسمى بالخطأ العشوائي. ولنفرض أنه في دراسة ما كان هناك متغير مستقل واحد (معالجة) ومتغير تابع واحد ففي هذه الحالة يقسم التباين الكلي إلى تباين ناتج عن أثر المعالجة أو تباين مفسر، بينما يعود باقي التباين إلى ما يسمى بالخطأ وفي هذه الحالة يمكن صياغة معادلة مربع إيتا على النحو التالي:

$$\text{مؤشر مربع إيتا} = \frac{\text{م م ع}}{\text{م م ك}}$$

حيث إن (م م ع) : مجموع مربع الانحرافات الناتج عن أثر المعالجة ، (م م ك) : مجموع مربع الانحرافات الكلي.

ولو فرضنا أن (م م ع) = ٣٥١,٥٢ وأن (م م ك) = ٧٨٦,٨ فإنه يمكن حساب قيمة مربع إيتا على النحو التالي:

$$\text{مؤشر مربع إيتا} = \frac{٣٥١,٥٢}{٧٨٦,٨} = ٤,٤٤$$

وهذا يعني أن ما نسبته ٤٪ من التباين الكلي يمكن تفسيره من خلال المعالجة. ت - مؤشر مربع أو ميجا (٤) : يعد مؤشر مربع أو ميجا أقل تحيزاً من مؤشر مربع إيتا ويمكن حسابه وفق المعادلة التالية:

$$\text{مربع أو ميجا} = \frac{\text{م م ع} - (\text{ج} - ١) \cdot \text{م م خ}}{\text{م م ع} + \text{م م خ}}$$

حيث إن (م م ع) : مجموع مربع الانحرافات الناتج عن أثر المعالجة. ، (ج) : عدد المجموعات أو عدد مستويات المتغير المستقل ، (م م خ) : متوسط مربع الانحرافات للخطأ ، (م م ك) : مجموع مربع الانحرافات الكلي.

ففي المثال ورد عند حساب قيمة مربع إيتا إذا افترضنا أن عدد المجموعات يساوي (٤) وأن متوسط مربع الإنحرافات للخطأ يساوي (٩,٦٧) وفي هذه الحالة يمكن حساب قيمة مربع أو ميجا على النحو التالي:

$$\text{مربع أو ميجا} = \frac{٩,٦٧ - (١ - ٤) \cdot ٣٥١,٥٢}{٩,٦٧ + ٧٨٦,٨} = ٠,٤٠$$

أي أنه وفق مؤشر مربع أو ميجا فإن ما نسبته ٤٪ من التباين الكلي يمكن تفسيره من خلال المعالجة ويلاحظ أن قيمة مربع أو ميجا أقل من قيمة مربع إيتا وهذه نتيجة متوقعة حيث إن مؤشر مربع إيتا - وكما ذكر سابقاً - يضم من قيمة التباين المفسر، ومن الجدير ذكره أن المعادلة الواردة هنا لمربع أو ميجا لا تصلح في حالة تصاميم القياسات المتكررة (Repeated-measures designs).

وفيما يتعلق بالحكم على مقدار مؤشر مربع إيتا أشار كوهين (Cohen, 1997) إلى

اعتبار هذا المقدار صغيراً عندما يساوي (٠,٠١) وإلى اعتباره متوسطاً عندما يساوي (٠,٠٦) وإلى اعتباره كبيراً عندما يساوي (٠,١٤). أما بالنسبة للحكم على مقدار مربع أو ميغا فإن الدراسات -حسب حدود علم الباحث- لم تشر إلى معايير محددة للحكم عليه ولعل ذلك مرده إلى أن مربع إيتا أكثر استخداماً من مؤشر أو ميغا وربما يعود ذلك إلى أن عملية حسابه أكثر سهولة. لكن نتيجة للتتشابه بين مؤشر مربع إيتا ومؤشر مربع أو ميغا من حيث طريقة التفسير ونتيجة لتقريب قيمة المؤشرين وخاصة مع العينات الكبيرة (التي تزيد عن ١٠٠) (Snyder & Lawson, 1993) فإنه يمكن استخدام نفس المعايير التي قدمها كوهين في حالة مربع إيتا للحكم على مقدار قيمة مربع أو ميغا.

٣- تقدير حجم الأثر في الدراسات الارتباطية:

على الرغم من وجود أنواع مختلفة من عاملات الارتباط -نتيجة اختلاف نوع معامل الارتباط المستخدم تبعاً لطبيعة البيانات- يعد معامل ارتباط بيرسون أكثر عاملات الارتباط استخداماً في الدراسات التربوية والنفسية ويستخدم لفحص مدى العلاقة الخطية بين متغيرين متصلين. ومن المعروف أن معامل ارتباط بيرسون يتراوح ما بين -١ إلى ١ وأن الإشارة التي تسبق قيمة معامل الارتباط إنما تدل على اتجاه العلاقة وليس على قوة العلاقة. ومن الملاحظ أن الدراسات العربية التي تستخدم معامل ارتباط بيرسون تشير فقط إلى مقدار واتجاه هذا المعامل دون الإشارة إلى مقدار التباين الذي يفسره كل متغير من تباين المتغير الآخر، أي دون الإشارة إلى مقدار مربع معامل الارتباط، والذي ما هو إلا أحد المؤشرات التي تدل على مقدار ترابط تباين كلا المتغيرين قيد الدراسة أو مقدار تأثير تباين كل من المتغيرين بتباين المتغير الآخر، وهذه هي نفس فكرة حجم الأثر، من هنا يمكن القول أن مربع معامل ارتباط بيرسون ما هو إلا أحد مؤشرات حجم الأثر.

٤- تقدير حجم الأثر في الدراسات التي تستخدم "الانحدار":

من المعروف أن هناك ثلاثة أنواع من الانحدار، وهي الانحدار البسيط (Simple regression) والانحدار المتعدد (Multiple regression) والانحدار متعدد التباين (Multivariate regression) حيث يعتمد النوع الأول على الارتباط البسيط (معامل ارتباط بيرسون) بينما يعتمد النوع الثاني على ما يسمى بالارتباط المتعدد (Multipile R) في حين يعتمد النوع الثالث على ما يسمى بالارتباط القانوني (Canonical correlation). وسيكتفى في البحث الحالي الإشارة إلى المؤشرات، التي تستخدم لتقدير حجم الأثر في النوعين الأول والثاني دون التطرق إلى كيفية استخراج هذه المؤشرات لأن ذلك يقع خارج إطار أهداف هذا البحث. في حالة الإنحدار البسيط يتم استخدام مقدار الارتباط بين المتغير المتنبأ والمتغير المتنبأ به لمعرفة مدى قدرة المتنبأ على تفسير تباين المتنبأ به من خلال تربيع قيمة معامل الارتباط بين المتغيرين. ويشار إلى هذه القيمة

مربع (ر) (R^2). أما بالنسبة لللانحدار المتعدد فهو يعتمد أيضاً على مربع (ر) ولكن للارتباط بين مجموعة من المتغيرات أو المتباينات من جهة ومتغير متباين به من جهة أخرى. وهناك معادلات خاصة لتقدير قيمة مربع (ر) سواء من خلال عينة الدراسة أو مجتمع الدراسة أو في مجتمعات وعينات جديدة. وينصح الراغب في الاستزادة حول هذا الموضوع مراجعة ستيفينز (Stevens, 1996).

العلاقة بين حجم الأثر وقوة الاختبار المتوقعه وحجم العينة.

بدايةً ومن المهم هنا الإشارة إلى أن هناك معادلات رياضية لتقدير قوة كل اختبار إحصائي على حدة (Howell, 1995) وإن عرض هذه المعادلات خارج أهداف البحث الحالي لكن من المهم أيضاً الإشارة إلى أن تقدير قوة أي اختبار إحصائي يعتمد على ثلاثة عناصر هي : حجم الأثر المتوقع وحجم العينة ومستوى الدلالة. بالنسبة لحجم الأثر المتوقع فإن المقصود به توقع مدى العلاقة بين المتغير المستقل أو المتغيرات المستقلة والمتغير أو المتغيرات التابعه ويمكن تقدير حجم الأثر المتوقع من خلال ثلاث طرق (Howell, 1995):

١- مراجعة الأدب المتخصص: من خلال مراجعة الدراسات السابقة التي تمت حول موضوع البحث ومن خلال عرض نتائج تلك الأبحاث يستطيع الباحث بناء توقع حول مقدار حجم الأثر خاصة إذا أشارت تلك الدراسات مباشرة إلى قيمة حجم الأثر التي ظهرت من خلال بياناتها. فعلى سبيل المثال إذا استخدم باحث اختبار "ت" لفحص مدى تأثير متغير مستقل معين على متغير تابع، فإن بإمكانه بالإضافة إلى عرض نتيجة ذلك الاختبار أن يفحص مقدار حجم الأثر باستخدام أحد المؤشرات الخاصة بذلك مثل مؤشر "د" أو جاما أو دلتا، وهو بذلك يقدم أساساً لغيره من الباحثين يمكنهم من بناء توقع معين حول العلاقة بين متغيرات الدراسة أو الدراسات التي هم بصدده إجرائها. ومن هنا جاءت أهمية استخراج قيمة حجم الأثر في الدراسات التربوية والنفسية حيث إن هذا الإجراء يخدم الأبحاث نفسها من حيث عدم الاكتفاء فقط برفض الفرضية الصفرية أو الفشل في رفضها بل أيضاً بمعرفة وتحديد حجم تأثير المتغير المستقل على المتغير التابع، أو معرفة مدى قدرة المتغير المستقل على تفسير تباين المتغير التابع؛ لذلك يمكن عد حجم الأثر أسلوباً إحصائياً مكملاً لفحص الفرضيات الإحصائية. من جهة أخرى فإن الإشارة مقدار حجم الأثر في الدراسات المختلفة تمكن الباحثين في المستقبل من صياغة توقع حول العلاقة بين متغيرات دراستهم.

٢- الطريقة الثانية لتقدير حجم الأثر هي أن يحدد الباحث مقدار الفارق الدال عملياً -من وجهة نظر الباحث- بين متوسط المجموعة التجريبية والمجموعة الضابطة في حالة وجود عيتيتين أو الفارق الدال عملياً بين القياسين القبلي والبعدي في حالة اختبار "ت" للعينات المترابطة أو الفارق الدال عملياً بين متوسط العينة ومتوسط مجتمع معين في حالة اختبار

"ت" للعينة الواحدة أو مقدار الترابط الدال علماً بين تباين المتغير التابع في حالة تحليل التباين . فعلى سبيل المثال إذا علم باحث أن متوسط أداء طلاب الصف الأول الثانوي في منطقته التعليمية في مقرر الرياضيات يساوي (٧٠) فإذا قام باختيار (٥٠) طالباً عشوائياً من نفس المرحلة وطبق عليهم برنامجاً تعليمياً مكتفياً فإنه يستطيع أن يحدد مسبقاً أن الفارق الذي يسعى إليه بين متوسط عينته ومتوسط المجتمع هو ٨ وحدات صالح عينة الدراسة . ومن العوامل التي تلعب دوراً في تحديد الباحث لمقادير الأثر المرغوب به هي دعم نظرية معينة، أو إثبات فعالية برامج معينة في تحقيق أهداف محددة وكذلك كلفة هذه البرامج حيث كلما زادت الكلفة المادية للبرامج المقترحة من قبل الباحثين كلما زادت الحاجة إلى إثبات الأهمية أو الفائدة العملية لهذه البرامج.

٣- يستطيع الباحث استخدام المعايير التي قدمها كوهين لتقدير حجم الأثر ففي حالة استخدامه اختبار "ت" مثلاً من المعروف -كما ذكر سابقاً- أن كوهين وأشار إلى أن حجم الأثر يعد صغيراً عند (٥٠,٢٠) ومتوسطاً عند (٨٠,٥٠) وكبيراً عند (٨٠,٨٠) من هنا يستطيع الباحث أن يحدد أنه يسعى إلى حجم أثر متوسط أي (٥٠,٥٠) أو كبير أي (٨٠,٨٠) أو أكثر وهذا يتم في ضوء قوة النظرية التي يستند إليها البحث وقوة العلاقة بين متغيرات الدراسة. أما بالنسبة للعلاقة بين حجم العينة وقوة الاختبار المتوقعة فإنه من المعروف أنه كلما زاد حجم العينة كلما ازدادت احتمالية رفض الفرضية الصفرية؛ وذلك لأنه كلما كبر حجم العينة كلما قل الخطأ المعياري (ع ن) وكلما ازدادت قيمة الإحصائي بغض النظر عن التوزيع العيني لذلك الإحصائي.

أما بالنسبة للعلاقة بين مستوى الدلالة وقوة الاختبار المتوقعة فإنه كلما إزداد مستوى الدلالة أو الاحتمالية التي يسمح بها الباحث لنفسه للوقوع في الخطأ من النوع الأول كلما ازداد احتمال رفض الفرضية الصفرية وبالتالي ازدادت قوة الاختبار المتوقعة.

الاستنتاجات:

بصورة عامة يمكن القول أنه في معظم الأحيان يكمن هدف الباحثين في الوصول إلى علاقة دالة إحصائية؛ ذلك أن الباحث لا يبني فرضياته من الفراغ، بل من خلال مراجعة دقيقة للنظريات والدراسات التي تمت حول العلاقة بين متغيرات دراسته ، من هنا يسعى الباحث مسبقاً إلى معرفة مدى احتمالية وصوله إلى نتائج دالة إحصائية، أو يعني آخر مدى قوة الاختبار الإحصائي المستخدم في فحص الفرضية أو الفرضيات الصفرية في زيادة احتمالية رفض هذه الفرضية أو الفرضيات وهي خاطئة وهذا ما يطلق عليه "قوة الاختبار الإحصائي". ومن أهم العوامل التي قوم بدور كبير في تحديد قيمة "قوة الاختبار" حجم الأثر المتوقع وحجم العينة ومستوى الدلالة الذي سيستخدم في فحص الفرضيات الصفرية. ولقد بنى الافتراض الذي مفاده أن الباحثين - في معظم الأحيان - يسعون إلى رفض

الفرضية الصفرية على أساس أنهم يبدأون بتقديم المبررات النظرية ونتائج العديد من الدراسات السابقة التي يشير إلى وجود علاقة بين متغيرات دراساتهم مما يجعلهم يتوقعون الحصول على نتائج دالة إحصائية. قد يقال بأن بعض الأبحاث تجرى أحياناً في ضوء تعارض نتائج الدراسات السابقة حيث يشير بعضها إلى وجود نتائج دالة إحصائية، بينما يشير بعضاً آخر إلى عدم وجود نتائج دالة إحصائية. ويقى السؤال الذي يبقى بحاجة إلى فحص وإجابة هو: إلى أي درجة يمكن الوثوق بإجراءات الدراسات التي توصلت إلى نتائج غير دالة إحصائية؟ وهل كانت هذه الإجراءات مصممة بصورة دقيقة بحيث تمكّن الباحثين من اكتشاف العلاقة بين متغيرات دراساتهم؟. معنى إلى أي درجة كان إجراء المتغير المستقل دقيقاً؟ هل تم ضبط المتغيرات الداخلية بحيث تم تحبيدها أو إدخالها في تصميم البحث؟ هل كان حجم العينة مناسباً؟ هل كانت أدوات القياس المستخدمة دقيقة؟ وفي هذا الصدد يشير بعض المتخصصين في منهج البحث أن الافتراض الأساسي في الدراسة التجريبية هو وجود فروق دالة إحصائية بين مجموعات الدراسة، وأنه إذا ما تم إجراء البحث وفق المنهجية السليمة فإنه لا بد عند استخدام الأسلوب الإحصائي المناسب - من رصد هذه الفروق. ويسمى المتخصصون في مجال البحث احتمالية رصد الأسلوب الإحصائي لهذه الفروق بقوة الاختبار المتوقعة. لكن من المهم أيضاً أنه عند الحصول على نتائج دالة إحصائية الإشارة إلى مقدار قوّة العلاقة بين متغيرات دراسته. أي إلى مقدار حجم أثر المتغير أو المتغيرات المستقلة على المتغير أو المتغيرات التابع. وأن تقرير مقدار هذا الأثر يساعد غيره من الباحثين على اختيار حجم العينة ومستوى الدلالة المناسبين لزيادة احتمالية قوة الاختبار الإحصائي المتوقعة. فعلى سبيل المثال إذا رغب الباحث المستخدم لاختبار "ت" للوصول إلى قوّة اختبار متوقعة مقدارها (٨٠، ٨٠) فإنه بحاجة أولاً إلى مراجعة الدراسات السابقة أو اتباع إحدى الطرق التي تم ذكرها لتحديد "حجم الأثر المتوقع" وفي ضوء ذلك يحدد حجم العينة المطلوب لتحقيق هذا الغرض.

ومن الواضح من خلال ما سبق أن فحص الباحث لمقدار حجم الأثر المستخرج من خلال بيانات دراسته يقدم فائدتين كثیرتين أولاهما أنه يقدم تقريراً حول مدى العلاقة بين المتغير المستقل والمتغير التابع، أي أنه لا يكتفي بالإجابة عن الفرضية الصفرية "نعم" أو "لا" بل إنه يحدد مقدار أثر المتغير المستقل على المتغير التابع حيث إنه في بعض الأحيان قد يكون الفرق بين المجموعة التجريبية والمجموعة الضابطة كبيراً ولكنه غير دال إحصائياً بسبب أحد العوامل التي تم ذكرها سابقاً. ومن جهة أخرى قد يكون هذا الفرق صغير جداً ولكنه دال بسبب كبير حجم العينة أو نتيجة استخدام مستوى دلالة مرتفع. أما الفائدة الثانية لتقرير "مقدار حجم الأثر" فتتمثل في أنه يزود الباحثين الآخرين بفكرة عن قوّة العلاقة أو حجم تأثير المتغير المستقل على المتغير التابع ، من هنا فإن حساب مقدار حجم الأثر يقدم معلومات وفائدة لا تقل أهمية عن عملية فحص الفرضيات الإحصائية، بل إنه يفصل ويووضح ويكمّل المعلومات المستخلصة من خلال فحص الفرضيات الإحصائية. وبناءً على ما سبق يوصي

الباحث زملاءه من الباحثين عومماً، والباحثين العرب على وجه الخصوص بالإشارة إلى مقدار حجم الأثر في دراساتهم وذلك لفهم نتائج دراساتهم بصورة أفضل واستخدام هذه النتائج فيما بعد لبناء أبحاث وفق أسس نظرية وتطبيقية سليمة.

المراجع

عوده ، أحمد و الخليلي ، خليل. (١٩٨٨). **الإحصاء للباحث في التربية والعلوم الإنسانية**. عمان –الأردن: دار الفكر للنشر والتوزيع.

- Arnemnn, K.(2003). **A Review of he panoply of effect size choices**. (ERIC Document Reproduction Service No. ED 473811).
- Aron, A., & Aron, E.N. (1997). **Statistics for the behavioral and social sciences: A brief course**. NJ: Prentice Hall.
- Cohen, J. (1977). **Statistical power analysis for the behavioral sciences**, New York: Academic Press.
- Cohen, J. (1988). **Statistical power analysis for the behavioral sciences**. Hillsdale, NJ: Erlbaum.
- Craig, J.R., Eison , C.L., & Metz , L.P. (1976). Significance tests and their interpretation: An example utilizing published research and Omega Squared. **Bulletin of the Psychonomic Society** , 7, 280-282
- Fan, Xitao.(2001).Statistical significance and effect size in education research:Two sides of a coin. **Journal of Educational Research**, 94, (5) 257-282.
- Fern, E.,Monroe,K.B. (1996).Effect-size estimates: Issues and problems in interpretation. **Journal of Consumer Research**, 23, 89-104.
- Hays, W.L. (1963). **Statistics for Psychologists**.New York: Holt , Rinehart , and Winston.
- Howell, D.C. (1995). **Fundamental statistics for the behavioral sciences**.(3rd ed.). CA: Belmont.
- Lane, G.(1999).**Show me the magnitude! the consequences of overemphasis on null hypothesis significance testing**. Paper presented at the annual meeting of the Mid-South Educational Research Association(AI, November 16-19?1999).(Eric Document Reproduction Service No. ED 436557).

- Mahadevan, L.(2000).**The effect size statistic: Overview of various choices.**(Eric Document Reproduction Service No. ED 438308).
- Maxwell, S.E., & Delaney , H.D. (1990). **Designing experiments and analyzing data: A model comparison perspective.** Belmont , CA: Wadsworth Publishing.
- McClain, A.(1995).**Effect size as an alternative to statistical significance testing.** Paper presented at the annual meeting for the American Educational Research Association. San Francisco, CA.(Eric Document Reproduction Service No. ED382639).
- McLean, J.E, Eronest,J.M.(1997). **Has testing for statistical significance outlived its usefulness.**Paper presented at the the annual meeting of the Mid-south Educational Research Association (26, Memphis, TN, November 12-14, 1997). (Eric Reproduction Service No.Ed415265).
- McMillan, J.H., Lawson, S., Lewis, K., & Snyder, A.(2002). **Reporting effect size: The road less traveled.** (ERIC Reproduction Service No.Ed466689).
- Schneider, A.L., & Darcy , R.E. (1984). Policy implications of using significance tests in evaluation research. **Evaluation Review**, *8*(4) , 573-583.
- Snyder, P., & Lawson , S. (1993). Evaluating results using corrected and uncorrected effect size estimates. **Journal of Experimental Education** ,*61*, 334- 349.
- Stevens, J.P. (1996). **Applied multivariate statistics for the social sciences** (3rd ed.). NJ: Lawrence Erlbaum.
- Tapia, M., & Marsh, G. (2002).**Interpreting and reporting effect sizes in research investigations.**(ERIC Document Reproduction Service No. ED 471293).
- Thompson, B. (1998). **Encouraging effect size reporting is not working: The etiology of research resistance to changing practices.** Paper presented at the annual meeting of the Southwest Educational Research Association (Houston , Tx , January 1998). (Eric Document Reproduction Service No. ED 416214).

